Research Note 2012-03

Notional Army Enlisted Assessment Program: Cost Analysis and Summary

Deirdre J. Knapp and Roy C. Campbell (Editors)
Human Resources Research Organization

Personnel Assessment Research Unit
Michael G. Rumsey, Chief

December 2011

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

MICHELLE SAMS Ph.D.
Director

Research accomplished under contract
for the Department of the Army

Human Resources Research Organization

Technical reviews by

Kimberly Owens, U.S. Army Research Institute

NOTICES

DISTRIBUTION: This Research Note has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This Research Note may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Research Note are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
December 2011	Final	January 2005 – January 2006

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Notional Army Enlisted Assessment Program: Cost Analysis and Summary	DASW01-03-D-0015/DO 0013
	5b. PROGRAM ELEMENT NUMBER 622785

6. AUTHOR(S)	5c. PROJECT NUMBER
Deirdre J. Knapp and Roy C. Campbell (Editors) (Human Resources Research Organization)	A790
	5d. TASK NUMBER 104
	5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Human Resources Research Organization 66 Canal Center Plaza, Suite 400 Alexandria, VA 22314	FR-05-75

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM
U. S. Research Institute for the Behavioral & Social Sciences ATTN: DAPE-ARI-RS 2511 Jefferson Davis Highway Arlington, VA 22202-3926	ARI
	11. MONITOR REPORT NUMBER Research Note 2012-03

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Contracting Officer's Representatives: Tonia Heffner and Peter Greenston

14. ABSTRACT *(Maximum 200 words)*:

In the early 1990s, the Department of the Army abandoned its Skill Qualification Test (SQT) program due primarily to maintenance, development, and administration costs. This left a void in the Army's capabilities for assessing job performance qualification. To meet this need, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) instituted a 3-year program of feasibility research related to development of a Soldier assessment system that is both effective and affordable. The PerformM21 program has had two mutually supporting tracks. The first track has focused on the design of a testing program and identification of issues related to its implementation. The second has been a demonstration of concept—starting with a prototype core assessment targeted to all Soldiers eligible for promotion to Sergeant, followed by job-specific prototype assessments for several Military Occupational Specialties (MOS). The prototype assessments were developed during the first 2 years of the research program. Pilot testing of the prototype assessments was completed in the third year of the project and is documented in a companion report. The present report describes the notional test program and analyzes the anticipated costs and describes the benefits associated with its implementation.

15. SUBJECT TERMS

behavioral and social science, personnel , job performance measurement, manpower, competency assessment

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	43	Ellen Kinzer Technical Publication Specialist (703) 545-4225

Standard Form 298

Acknowledgements

U.S. Army Research Institute for the Behavioral and Social Sciences (ARI)
Contracting Officer Representatives (COR)

Dr. Tonia Heffner and Dr. Peter Greenston of ARI served as co-COR for this project, but their involvement and participation went far beyond the usual COR requirements. Their contributions and active input played a significant role in the production of the final product and they share credit for the outcome. Of particular note are their activities in conveying information about the project in briefings and presentations to Army Leadership on many important levels.

The Army Test Program Advisory Team (ATPAT)

The functions and contributions of the ATPAT, as a group, are documented in this report. But this does not fully reflect the individual efforts that were put forth by members of this group. Project staff is particularly indebted to Sergeant Major Michael Lamb, currently with Army G-3, who served as the ATPAT Chairperson during this work.

The other individual members of the ATPAT who were active and involved during this phase were:

SGM John Cross
CSM George D. DeSario
SGM (R) Julian Edmondson
CSM Dan Elder
CSM (R) Victor Gomez
SGM John Griffin
SGM John Heinrichs
SGM (R) James Herrell
SGM Enrique Hoyos
CSM Nick Piacentini
SGM David Litteral
SGM Michael Magee
SGM Tony McGee
SGM John Mayo
SGM Pamela Neal
CSM Doug Piltz
SGM (R) Gerald Purcell
CSM Robie Roberson
CSM Otis Smith Jr
MSG Matt Northen

NOTIONAL ARMY ENLISTED ASSESSMENT PROGRAM: COST ANALYSIS AND SUMMARY

EXECUTIVE SUMMARY

Research Requirement:

The Army Training and Leader Development Panel NCO survey (Department of the Army, 2002) called for objective performance assessment and self-assessment of Soldier technical and leadership skills to meet emerging and divergent Future Force requirements. The Department of the Army's previous experiences with job skill assessments in the form of Skill Qualification Tests (SQT) and Skill Development Tests (SDT) were reasonably effective from a measurement aspect but were burdened with excessive manpower and financial resource requirements.

Procedure:

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) conducted a 3-year feasibility effort to identify viable approaches for the development of a useful yet affordable operational performance assessment system for Army enlisted personnel. Such a system would depend on technological advances in analysis, test development, and test administration that were unavailable in the previous SQT/SDT incarnations.

The ARI project (known as PerformM21) was conducted with support from the Human Resources Research Organization (HumRRO) and entailed three phases:

- Phase I: Identify User Requirements, Feasibility Issues, and Alternative Designs
- Phase II: Develop and Pilot Test Prototype Measures
- Phase III: Evaluate Performance Measures, Conduct a Cost-Benefit Analysis, and Make System Recommendations

The objective of Phase I was to identify issues that the overall recommendation needs to take into account for a viable, Army-wide system (Knapp & Campbell, 2004). Phase I also produced a rapid prototype assessment covering Army-wide "core content" with associated test delivery and test preparation materials (Keenan, Campbell, Moriarty, Knapp, & Heffner, 2006).

In Phase II, the research team (a) pilot tested the core competency assessment, (b) developed competency assessment prototypes for five Military Occupational Specialties (MOS), and (c) explored issues further to develop more detailed recommendations related to the design and feasibility of a new Army enlisted personnel competency assessment program. The work in Phase II is documented in Knapp and Campbell (2005).

In Phase III, the MOS tests (along with the common core examination) were pilot tested and a cost and benefit analysis of a notional Army program was conducted. Because it was not possible to derive defensible dollar estimates associated with anticipated program benefits, we articulated the benefits as part of this analysis, but did not quantify them. The cost and benefit analysis, along with recommendations related to the notional assessment program are presented in this report. The Phase III pilot test activities are documented in a companion research report (Moriarty & Knapp, 2007).

Findings:

The main conclusion from the PerformM21 work is that a testing program that includes just an Army-wide "core competency" assessment is quite feasible and likely to be cost-effective. Introducing MOS-specific testing will substantially increase costs, particularly if we assume that most MOS would have their own tests and that these tests would include some relatively expensive measurement methods (e.g., hands-on tests, computer-based simulations). If the Army views these costs as excessive, it would also be reasonable to consider a somewhat scaled back program that would not include all MOS and/or excludes some of the higher cost assessment methods.

Utilization of Findings:

The program design and technology issues and recommendations resulting from this feasibility research are intended to help Army leaders make informed decisions about the future of competency assessment for the enlisted force. The parallel prototyping work has resulted in lessons learned and test content suitable for incorporation into an operational test program.

NOTIONAL ARMY ENLISTED ASSESSMENT PROGRAM: COST ANALYSIS AND SUMMARY

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Preface

This report describes the cost analysis of a notional assessment program that grew out of the PerformM21 research program conducted over the 2003–2006 time period. The objective of this effort was to develop cost-effective measures that realistically tracked demands of performance on the job. Our conclusion is that the objective was met in the sense that advances were made both in the development of quality measures and in the identification of techniques that minimized the cost of such measures.

With respect to the Army-wide measures, the costs appear manageable. At the present time, however, one can question whether the implementation of such tests on an MOS (military occupational specialty) by MOS basis is feasible. The benefits were not quantified here, so it is difficult to say to what extent the benefits are commensurate with the costs. In the current, extremely resource-constrained, environment, it might well be supposed that funding for the costs identified here, despite their reasonableness when one considers the many jobs that need to be covered, would be difficult to obtain. Since there has been no implementation of these tests since this project was completed, that would indeed seem to be the case.

However, it would be unfortunate if this analysis provided the basis for the conclusion that any kind of performance testing in the Army is infeasible. One constraint we are currently operating under is the lack of a complete understanding of the degree to which different aspects of a job or task need to be represented in order to have a test which truly reflects an individual's capability to perform a particular job. If we could identify underlying competencies that were sufficiently generalizable across tasks such that separate measures of each related task were not needed, the costs associated with performance test development could be dramatically reduced. Methods for identifying such competencies have been advanced in the past, but it is our sense that the competencies identified using such methods were too general to properly represent the associated tasks. At the present time we are engaged in research that may hopefully lead to a more favorable outcome. It is our intention to continue to explore means whereby the costs of job analysis and performance testing can be reduced to the point where the advantages of implementing such tests are incontrovertible.

Michael Rumsey & Peter Greenston
November 2011

NOTIONAL ARMY ENLISTED ASSESSMENT PROGRAM: COST ANALYSIS AND SUMMARY

CHAPTER 1: PERFORMM21 RESEARCH PROGRAM OVERVIEW

Deirdre J. Knapp and Roy C. Campbell (HumRRO)

Introduction

Individual Soldier readiness is the foundation of a successful force. In the interest of promoting individual Soldier performance, the U.S. Department of the Army has previously had assessment programs to measure Soldier knowledge and skill. The last incarnation of such a program was the Skill Qualification Test (SQT) program. The SQT program devolved over a number of years, however, and in the early 1990s the Army abandoned it entirely due primarily to maintenance, development, and administration costs.

Cancellation of the SQT program left a void in the Army's capabilities for assessing job performance qualification. This was illustrated most prominently in June 2000, when the Chief of Staff of the Army established the Army Training and Leader Development Panel (ATLDP) to chart the future needs and requirements of the Noncommissioned Officer (NCO) corps. After a 2-year study, which incorporated the input of 35,000 NCOs and leaders, a major conclusion and recommendation was that the Army should: "Develop and sustain a competency assessment program for evaluating Soldiers' technical and tactical proficiency in the military occupational specialty (MOS) and leadership skills for their rank" (Department of the Army, 2002).

The impetus to include individual Soldier assessment research in the U.S Army Research Institute for the Behavioral and Social Sciences' (ARI's) programmed requirements began prior to 2000 and was based on a number of considerations regarding requirements in Soldier selection, classification, and qualifications. For example, lack of operational criterion measures has limited improvements in selection and classification systems. Meanwhile, there were several significant events within the Army that reinforced the need for efforts in this area. The aforementioned ATLDP recommendation resulted in the Office of the Sergeant Major of the Army (SMA) and the U.S. Army Training and Doctrine Command (TRADOC) initiating a series of reviews and consensus meetings with the purpose of instituting a Soldier competency assessment test. Ongoing efforts within the Army G-1 to revise the semi-centralized promotion system (which promotes Soldiers to the grades of E5 and E6) also were investigating the use of performance (test)-based measures to supplement the administrative criteria used to determine promotion. Ultimately, the three interests (ARI, SMA/TRADOC, G-1) coalesced and the ARI project sought to incorporate the program goals and operational concerns of all of the Army stakeholders, while still operating within its research-mandated orientation.

To meet the Army's need for job-based performance measures, ARI instituted a 3-year program of feasibility research, *Performance Measures for the 21st Century (PerformM21)*, to identify viable approaches for development of a Soldier assessment system that is both effective and affordable. This research has been conducted with contract support from the Human

Resources Research Organization (HumRRO) and its subcontractors, Job Performance Systems, Inc, The Lewin Group, and the SAG Corporation.

Research Program Overview

The PerformM21 research program is best viewed as having two mutually supporting tracks. The first track is essentially the conceptualization and capture of issues, features, and capabilities related to an Army-wide testing program. The second track is to develop and administer prototype tests and associated tools. These prototypes include both an Army-wide "common core" assessment and some selected MOS tests. These are intended to reflect, inasmuch as possible, design recommendations for the future operational assessment program. Experiences with the prototypes, in turn, influenced elaboration and modification of the operational program design recommendations as they developed during the course of the 3-year research program.

Formally, PerformM21 has had three phases:

- Phase I: Identify User Requirements, Feasibility Issues, and Alternative Designs
- Phase II: Develop and Pilot Test Prototype Measures
- Phase III: Evaluate Performance Measures, Conduct a Cost-Benefit Analysis, and Make System Recommendations

Phase I of PerformM21 resulted in program design recommendations that included such considerations as how an Army assessment would be delivered, how assessments would be designed, developed, and maintained, and what type of feedback would be given. It is at this point that certain basic assumptions were made that helped drive the remainder of the project. These included the assumption that the scores on the new Army tests would eventually be used as a consideration in promotion decisions and would thus require a high stakes testing model (e.g., proctored testing).

In Phase I, we also developed a demonstration common core assessment test to serve as a prototype for the envisioned new Army testing program. This core assessment is a computer-based, objective test that covers core knowledge areas applicable to Soldiers in all MOS (training, leadership, common tasks, history/values). Phase I was completed in January 2004 and is documented in two ARI publications (Knapp & Campbell, 2004; Keenan, Campbell, Moriarty, Knapp, & Heffner, 2006).

Phase II of the PerformM21 program (which corresponds roughly to year two of the 3-year overall effort) had three primary goals:

- Conduct an operational pilot test of the common core assessment with approximately 600 Soldiers.
- Investigate job-specific competency assessments. This resulted in prototype assessments for five MOS.

- Continue to refine and to develop discussion and recommendations related to the design and feasibility issues established in Phase I.

This work is detailed in an ARI technical report edited by Knapp and Campbell (2005).

The primary activities in Phase III were to (a) pilot test the prototype MOS-specific assessments (as well as conduct further pilot testing of the common core test), (b) conduct a cost-benefit analysis of the notional assessment program, and (c) make final recommendations. The pilot test work is detailed in a companion report (Moriarty & Knapp, 2007). A description of the notional assessment program and the cost-benefit analysis, as well as overall recommendations resulting from the entire PerformM21 3-year feasibility research effort, is provided in the present report.

Related Efforts

In addition to the core elements of PerformM21 broadly outlined in the three phases, there have been two related studies generated by requirements uncovered during the PerformM21 research. The first was a research effort to determine the kinds of information Soldiers need to determine their overall readiness for promotion, including identification of strengths and weaknesses prior to testing (Keenan & Campbell, 2005). This effort produced a prototype self-assessment tool intended to help prepare Soldiers for subsequent assessment on the common core test.

The second research effort was designed to determine new or refocused skills and tasks associated with operations in Iraq and Afghanistan and to include those requirements in a common core assessment program. The effort produced two major products. One was a prototype field survey designed to support development of a common core test "blueprint" and the second was development of additional common core test items targeted to content areas suggested by lessons learned in recent deployment operations. This work is documented in Moriarty, Knapp, and Campbell (2006).

The Army Test Program Advisory Team (ATPAT)

Early in Phase I, ARI constituted a group to advise us on the operational implications of Army assessment testing, primarily as part of the needs analysis aspect of the project. This group is called the Army Test Program Advisory Team (ATPAT) and the members are primarily Command Sergeants Major and Sergeants Major. ATPAT members represent key constituents representing various Army commands and all components. After the needs analysis, the ATPAT took on a role as oversight group for the common core and MOS assessments including serving as a resource for identifying and developing content for the tests. Eventually, the group became an all-around resource for all matters related to potential Army testing. The ATPAT also served as a conduit to explain and promote the PerformM21 project to various Army agencies and constituencies.

Research Approach: Integrating Process and Results

A key to organizing our approach has been the *Needs Analysis Organizing Structure.* Figure 1 lists the key components; the organizing structure is more fully explained in the Phase I needs analysis report (Knapp & Campbell, 2004). This structure helped organize our thinking and suggested the questions we posed to those providing input into the process. We obtained input from several sources as we considered the issues, ideas, and constraints associated with each requirement listed in Figure 1. These included the following:

- The Army Test Program Advisory Panel (ATPAT)
- Historical information about the SQT program and associated lessons learned
- Enlisted personnel promotion testing programs operated by the Air Force and the Navy
- Civilian assessment programs (e.g., professional certification and licensure programs)
- A review of automation and technology tools and systems

- Purpose/goals of the testing program
- Test content
- Test design
- Test development
- Test administration
- Interfacing with candidates
- Associated policies
- Links to Army systems
- Self-assessment

Figure 1. Outline of PerformM21 needs analysis organizing structure.

Purpose and Overview of Report

The purpose of this report is to abstract the major ideas, issues, and recommendations that have emerged from the 3-year PerformM21 feasibility research effort. Chapter 2 describes a notional test program that supports the program goals established at the start of the project, modified through experience gained during the course of the research. Given that the cost of a test program will be a major consideration in any implementation decision, Chapter 3 describes the process and results of a cost-benefit analysis effort. Chapter 4 provides an overall summary and discussion of the feasibility of implementing an assessment program. In the 3 years this research has been underway, ideas have surfaced about test programs that would have somewhat different goals than those on which the PerformM21 research was prefaced (e.g., dropping the link to promotion points). This last chapter, then, also discusses some of the implications of such shifts in focus for the design of an assessment program.

CHAPTER 2: A NOTIONAL ASSESSMENT PROGRAM

Deirdre J. Knapp (HumRRO)

Introduction

In prior PerformM21 project reports (Knapp & Campbell, 2004, 2005), we have offered recommendations and associated rationales for the design of a new Army assessment program. The purpose of this chapter is to provide a simple description of the envisioned program. It is this notional program that provided the basis for the cost-benefit analysis described in Chapter 3.

It is important to stress that many features of the notional test program would likely change as the Army moved forward with planning and implementation activities. Some deviations from the notional program (e.g., the size and make-up of an Army Assessment Office) would have little impact on its feasibility, costs, and benefits as outlined in this report. Other deviations (e.g., increased testing frequency) would more dramatically impact the program costs and outcomes. In fact, as of this writing, the Army is funding an effort to develop an initial test program (building off PerformM21 prototype tests) that would be used to support high priority MOS reclassification requirements. So long as such a program meets some minimal criteria (e.g., preserving the security of test item banks), it will further the Army's progress toward the broader assessment program and associated benefits described here.

Test Program Overview

Table 1 lists the major design features of the notional assessment program. These features have remained largely unchanged since the beginning of the PerformM21 feasibility research effort 3 years ago. An exception is that we originally planned to include E7 NCOs in the test population, but scaled the plan back to pay grades E4 through E6. This was done just prior to conducting the cost analysis work and reflected the collective judgment that the effort required to realistically include the associated costs for testing at the E7 pay grade outweighed the likelihood that the Army would include E7 NCOs in the assessment requirement, at least in the foreseeable future.

Table 1. Major Design Features of Notional Test Program

- All Soldiers in pay grades E4 through E6 will be included in the program
- Scores will be used to support promotion decisions
- The assessment program will be the same for all components of the Army
- There will be an Army-wide core competency test and/or MOS-specific tests
- Assessments will be computer delivered in a proctored environment
- Each test will be administered during a test window period each year
- Soldiers will be given adequate tools to prepare for the tests
- Scores will be valid for a 3-year period

The program's major design features are intended to maximize the positive impact of the program and strike a reasonable balance against program costs. For example, linking scores to promotion decisions will improve those decisions and help ensure that Soldiers are motivated to

prepare, thus increasing their job-relevant knowledge and overall readiness. Such high stakes testing requires a test program with security features that increase costs and reduces convenience (e.g., Soldiers cannot test anytime or anywhere). As another example, the notional program requires individual Soldiers to be tested once every 3 years rather than annually. While requiring Soldiers to test every year would help ensure that important job skills do not decay, it would significantly increase test program costs. The 3-year plan also fits into the Army's planned unit-focused stability program (now known as ARFORGEN).

The design features do not include a specific conclusion about whether the test program would include tests suitable for all Soldiers, regardless of MOS (i.e., an Army-wide "common core" assessment), MOS-specific tests, or both. An ideal program would likely eventually include both common core and MOS testing, so we have included both in our notional program. MOS testing greatly increases costs, however. Therefore, Chapter 3 will provide costs for common core testing both with and without MOS testing.

Process

In this section, we describe some of the mechanics behind the implementation of the notional assessment program. We have organized this discussion into the following areas:

- Policy, oversight, and coordination
- Assessment development and maintenance
- Assessment delivery

Policy, Oversight, and Coordination

The top part of Figure 2 shows that policy decisions, oversight, and program coordination would be achieved through the efforts of an overall program director, MOS directors (assuming MOS-specific testing), and a test council of senior NCOs. The bottom part of Figure 2 illustrates the functions that are required to support the Army testing program. All of this would be supported by a newly-established Army Assessment Office staffed by testing professionals, Army personnel, and administrative support persons. This office would also acquire, manage, and maintain the contractor and information technology (IT) systems needed to support the test program and carry out a variety of administrative functions. These include scheduling, records management, and communicating with Soldiers throughout all phases of the test process.

Figure 3 shows how the testing process would function as part of the Army promotion system. The Army Assessment Office will maintain databases of item-level data, final test scores, and associated information and transmit Soldier scores to the Army's central personnel database. This office will also provide score reports to Soldiers and rolled up score information to Army leaders at various levels. Total test scores will become part of Soldiers' records and enter into promotion decisions. For example, test scores would be integrated into semi-centralized system promotion point worksheets (e.g., on a 200-point scale). Soldiers will be given diagnostic feedback on their test performance (e.g., information about how they performed on different parts of the test). No pass/fail point will be established for the tests, since this would severely truncate and unnecessarily limit the informational value of the test program.

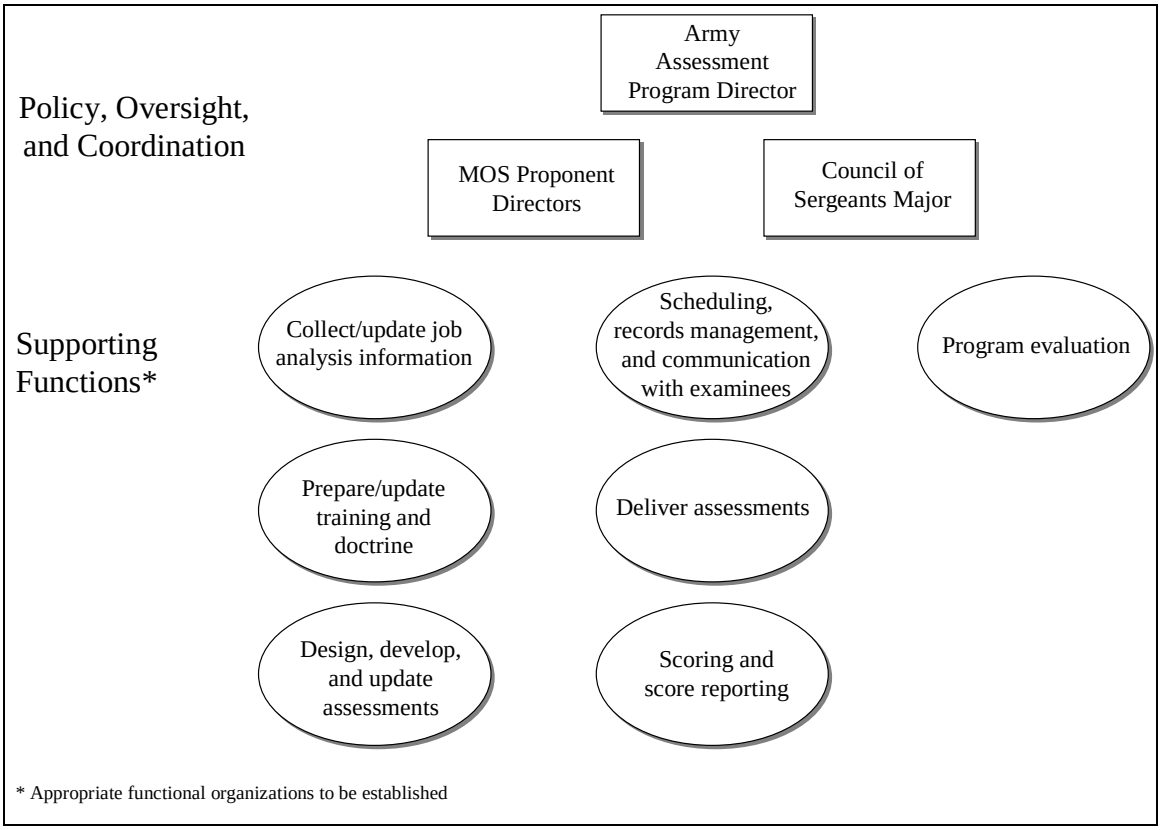

Figure 2. Assessment program supporting structure and functions.

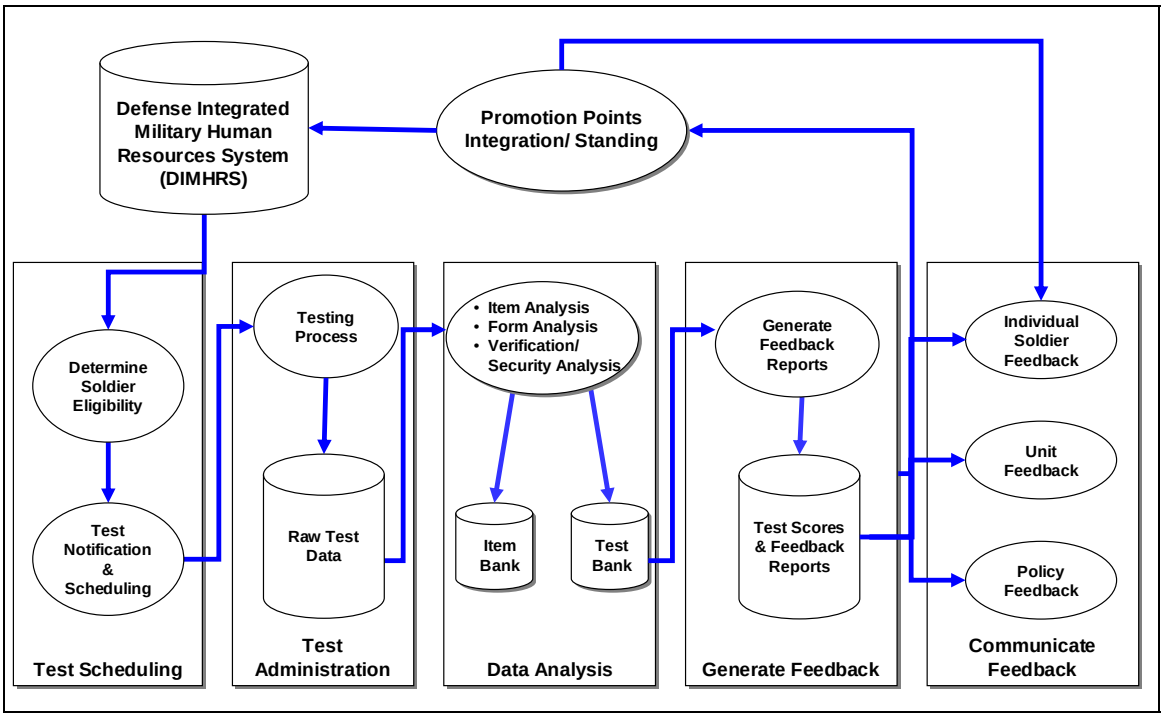

Figure 3. Test process flow diagram.

Assessment Development and Maintenance

Assessment Methods

Table 2 shows the test methods that might be used to assess Soldiers under this program. By far the most widely used method will be what we have variously called "enhanced multiple-choice (EMC)" or "job knowledge" tests. In both cases, the nomenclature is a bit misleading. Such tests (a) broadly cover relevant job/task knowledge, (b) include multiple-choice and other item formats (e.g., matching, ranking), and (c) make liberal use of graphics and animation to make the test experience more interesting and realistic and to reduce the reading requirement. Another fairly widely used method will be situational judgment tests (SJTs). This type of test presents problem scenarios drawn from actual Soldier experiences and lists several (usually four) actions a Soldier might take to respond. Examinees can either be asked to select the most and least effective actions or rate the effectiveness of each action. We used both strategies in the PerformM21 prototype tests.

Other methods include computer-based simulations that vary in content coverage and complexity. For example, a simple "path" simulation would require the Soldier to take actions throughout the simulated activity or task, but would keep the Soldier on the right track even if a prior action was incorrect. A more complex simulation (often called a multiple-path simulation) would respond to Soldier actions, making the assessment more realistic, but considerably more expensive to design and develop.

Table 2. Assessment Methods

Enhanced Multiple-Choice Tests (Job Knowledge Tests) ✓ Questions posed in applied work contexts ✓ Visual aids to reduce reading and enhance realism (e.g., photos, figures) ✓ Animation to enhance realism ✓ Non-traditional item formats (e.g., matching, drag-and-drop)
Situational Judgment Tests ✓ Real-life problem scenarios depicted in writing or through video ✓ Examinees evaluate effectiveness of various possible actions ✓ Focus is on judgment rather than knowledge, per se ✓ Scoring key based on expert judgment (e.g., senior NCOs)
Path Simulations ✓ Examinees are presented with a computer simulation of a problem scenario ✓ Examinees progress through the simulation, stopping at various points to answer questions
Complex Simulations ✓ Examinees are presented with a computer simulation of a problem scenario ✓ Examinees interact with the simulation, affecting how the scenario unfolds
Hands-On Tests ✓ Examinees perform job tasks in a standardized environment ✓ Performance is scored by expert observers

Hands-on tests are familiar to the Army. They are typically designed to cover specific tasks and are generally easy to develop because Army tasks tend to be highly proceduralized (i.e., there are explicit steps and standards associated with each task). The great expense with this method is associated with test delivery. A high stakes test model requires strict adherence to standardized test administration and scoring, which can be very difficult with hands-on tests that need to be delivered at many locations. We assume that hands-on tests would be administered at locations throughout the Army, with equipment, facilities, and supporting personnel provided by the host sites and traveling teams of test administrators/scorers provided by contractors.

In the notional assessment program, we assume that common core assessment will include an enhanced multiple-choice test and a situational judgment test, of which prototypes have been developed and administered in the PerformM21 research. We assume that all MOS that have MOS-specific tests would include an enhanced multiple-choice test and that many MOS would also include at least one additional assessment method (e.g., situational judgment test, simulation, or hands-on tests).

Designing Tests

In order to develop tests with job-relevant content, in-depth analysis of job requirements is needed. The notional test program relies on the Army's existing occupational analysis program (see Army Regulation 370-50) to provide a foundation for MOS-specific test content specifications. Under this program, MOS proponents are required to update occupational analysis information for their MOS every 3 years. The Army also periodically conducts occupational analysis aimed at identifying and updating common (i.e., Army-wide) task requirements. These programs are designed to provide information needed to update training aids (e.g., Soldier's Manuals) and training curricula, but can also provide a starting point for test design specifications.

Using the training-oriented occupational analysis information as a starting point, we have developed a prototype test design or "test emphasis" analysis process and associated Soldier survey (Moriarty et al., 2005). This survey would be administered at least every 3 years and would provide the information needed to update the test blueprint for each enhanced multiple-choice test. Test blueprints will detail categories of content to be covered by the test, including how much each category will be weighted on the test (e.g., 20% of the total test score will be based on first aid and 25% on weapons).

Job information needed to support the design and development of the other types of assessments varies by method. For example, simulations require identifying in detail the equipment, procedural steps, and contextual features of task performance. Often, collection of this information is integrated with the test development activities; however, in some situations the analysis might be specifically tailored to support these needs.

Developing Test Content

Test content (items, simulations) will be developed mostly by relying on contractors who employ psychometric professionals as well as former Army personnel with technical subject

matter expertise. These items will be reviewed for accuracy by active Army personnel. Over time, test "item banks" will grow and make development of the multiple equivalent test forms required for large-scale testing relatively easy. For the most part, it will be possible to pilot test new test items by embedding them on operational test forms, helping to ensure that only high quality items are used as a basis for scores reported back to Soldiers and their leadership.

Because higher skill level Soldiers are responsible for job content associated with lower skill levels, we anticipate considerable overlap in test content across pay grades. Thus, a single test item bank will be created for the core content and for any MOS tests. Items will be flagged to indicate to which pay grade they are applicable. It is even possible that, in some cases, Soldiers in different skill levels (e.g., E5 and E6) will receive the same test if the occupational analysis does not indicate differences that need to be reflected in the test.

Developing Test Forms

As discussed further in the next section, the notional test program calls for one test cycle per year for each MOS/pay grade (in which about one-third of Soldiers would test each year). For planning purposes, a test instrument will consist of about 100 items (or comparable measurement points, depending on the test method) that can be administered within a 2-hour testing period. Multiple equivalent test forms will be developed for each test as a security measure. Equivalent test forms will reflect the test blueprint, but each will have some reasonable percentage (e.g., 30%) of unique test content. Because content equivalent test forms might vary in difficulty, test score equivalence will be ensured through statistical methods (e.g., item response theory, equating based on an anchor test form).

Assessment Delivery

In our notional test program, computer-based tests will reside on a commercial server to be accessed by Soldiers scheduled for test sessions at Army Digital Training Facilities (DTFs) or National Guard Distributive Training Technology Project facilities (DTTPs). The commercial test delivery company will transmit test data to the Army Assessment Office for analysis and scoring by the applicable testing contractor.

Administering tests "on-demand" (i.e., at any time the test-taker desires) provides ultimate flexibility for test-takers, but unreasonable test development requirements for a high stakes testing program. This is because test content quickly becomes so widely shared that future test-takers can simply memorize material rather than learn it. On the other hand, the tightest control on test security—administering the test to everyone at one time—is impractical for today's Army. As with many civilian high stakes test programs, the notional Army program involves the use of multiple equivalent forms (the exact number based on the number of examinees) administered within an annual "test window." The length of the test window would be based on the volume of examinees, but would likely range from 2 to 4 months.

The notional program would start testing E4 Soldiers as soon as they become eligible for promotion to the E5 pay grade and Soldiers would continue to test until promoted to the E7 pay grade. Soldiers would be required to retest at 36 months although there would be a provision to

allow Soldiers who wanted to improve their scores to voluntarily retest annually. The latest test taken would become the score of record.

While it would be possible to administer the core examination and MOS tests at different points in time, it would be most efficient to administer them at the same time. If this were the case, then there would be a test window for E4 pay grade tests, another test window for E5 pay grade tests, and so forth. Rolling test windows have the advantage of spreading out the administrative effort to develop and administer tests rather than having a single high intensity administration period each year.

Hands-on tests would, of course, be administered apart from testing within the DTF/DTTP complex and likely on a unique schedule. Although hands-on tests would likely not be very widespread, they could affect some of the higher density MOS (such as infantry). We would therefore project a notional hands-on test program based on the unit location and schedule, much like the Expert Infantry Badge testing is currently being conducted. Contractor test teams would travel to locations and test throughout the year. Recording and transmittal of performance and scoring would be via hand-held computer technology.

Summary

The notional test program covers all aspects of a program from organizational oversight and policy setting, through assessment delivery and maintenance (including analysis), and finally through to assessment delivery. Our notional program is based on the testing design and functional requirements established through analysis of the needs of an Army test program. The program builds on many assumptions and ‘best idea” suppositions that may or may not be realized in an operational program. However, the notional program was an absolute requirement to facilitate the cost and benefits analysis, as presented in the next chapter.

CHAPTER 3: COST-BENEFIT ANALYSIS PROCEDURE AND RESULTS

Patrick Mackin and Kimberly Darling (SAG Corporation)
Carol Moore and Paul Hogan (The Lewin Group)

Introduction

One of the main goals of Phase III of Perform21 was to assess what an Army test program would cost. As can be seen from the description of the notional program described in Chapter 2, such a goal is challenging on many fronts. Foremost is that the program would not be isolated— it would have some involvement of many agencies and interests not the least of whom is the Soldier population in pay grades E4 though E6 who are the focus of the test program. The second challenge is that the program is embryonic—there is little existing structure, policy, or procedure currently existing within the Army on which to base solid projections. As a result, we made many assumptions about policy and practice that have yet to receive serious Army consideration or endorsement. Chapter 2 describes how an Army test program *could* work; not necessarily how it *will* work. But these assumptions are necessary in order to produce a workable cost model.

Using the best available data, we developed an activity-based costing model that will accommodate further refinement and expansion as additional data become available and the proposed program becomes better defined. In our model we present cost estimates for two variations as described in Chapter 2. The first scenario includes only an Army-wide test. The second scenario includes both Army-wide testing and MOS-specific testing.

We also address the issue of the benefits of the testing program. Most of the benefits deal with improved NCO selection and Soldier readiness and do not fit a quantifiable economic cost-benefit model. This does not make them any less real or desirable. Finally, despite limited data, we draw some cost comparisons between the notional PerformM21 program and the Army's previous SQT program.

Overview of the Cost Model

The cost model is activity based. That is, costs in the model are driven by a series of events related to the creation, delivery, and maintenance of the assessments. The main variables that drive costs include the number of assessments per year and the types of tests used.

Figure 4 illustrates the elements of the cost model. The main activity categories are (a) policy oversight and coordination, (b) assessment development, (c) assessment maintenance, and (d) assessment delivery. Through these activities flow the number of assessments—who is tested, the scope of testing (Army-wide and MOS scenarios), and the types of tests. These variables produce the cost projections.

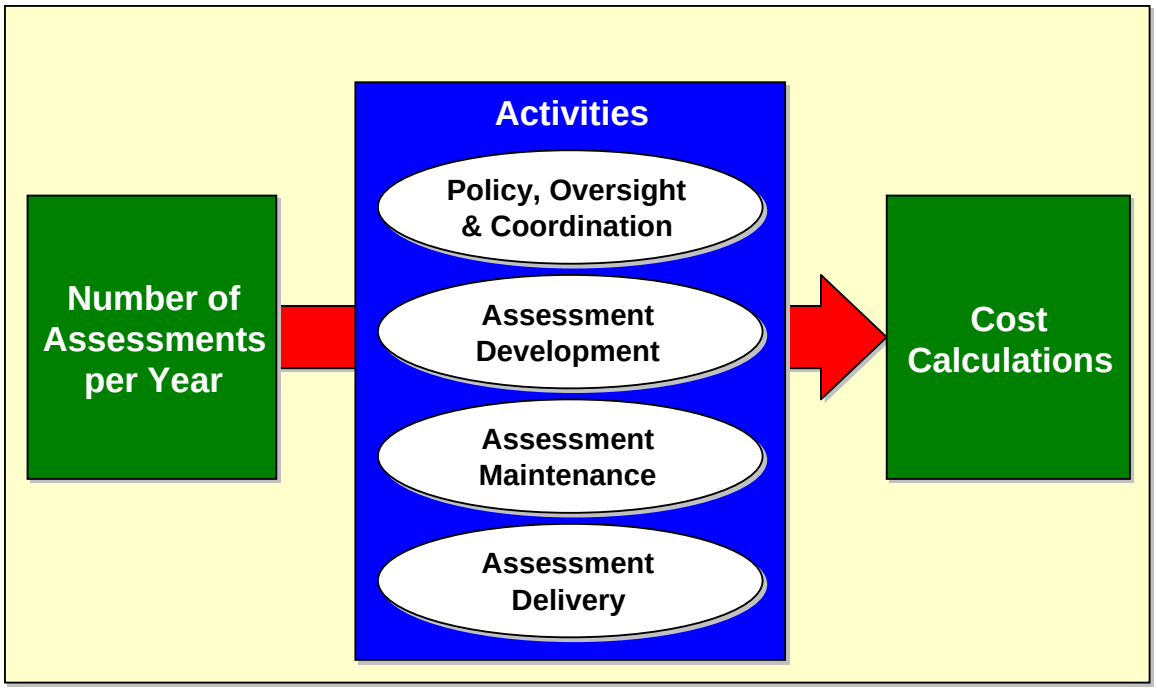

Figure 4. Cost model structure.

To forecast the costs of the assessment program, we estimated the resources that would be required to develop, maintain, and manage the program. We estimated costs for each major activity category described above, as well as the overhead costs associated with the Army Assessment Office. We also included Soldier costs for test preparation and test-taking. There are two main types of costs projected—the first are start-up or initial costs required to get a test program up and operating. These were projected over a 5-year time period. The second are annual operating costs, assuming a program already implemented.

Focusing on unit costs—rather than total costs—gave us the flexibility to scale budget projections to a variety of policy-driven assumptions, assuming linearity in the cost relationships. For example, the number of questions that would need to be included in a multiple-choice test bank is a function of the frequency with which Soldiers are tested and the length of testing windows. The total number of Soldiers tested has an indirect effect on the size of the test bank because it may constrain the lower bound of the testing window, given test facility capacity. The main resource requirements were professional labor (differentiated by experience level, subject matter expertise, and educational attainment), technical specialists, administrative support, Army subject matter experts (generally, retired Army personnel), travel, and investments in information technology (IT).

Sources of Cost Information

We collected data on the number and type of resources required via interviews with knowledgeable individuals from a variety of government and private organizations. We validated cost parameters gathered from interviews through the use of published averages and/or interviews with additional experts.

To identify interview participants who could provide relevant information about resource requirements for a particular activity, we first considered whether the Army would perform the activity itself or purchase services from the private sector. In accordance with the vision for the assessment program depicted in Chapter 2, we assumed that the Army would take advantage of expertise in the private sector for the bulk of the labor required in the test program. Expected contractor tasks include developing assessment instruments, maintaining test banks, and capturing test score data. Thus, our cost estimates for these activities are based on information gathered from experts in the private sector.

We assumed that functions generally recognized as governmental (e.g., oversight, policy, and limited direct Soldier participation) would be performed by the Army. For these, we interviewed officials from the relevant Army office regarding the type of labor required by GS-series and grade. For some services, such as test delivery, we explored the potential of Army, other government, and private sector providers to provide the service. In these cases, we based our projections on the lowest-cost combination of providers, consistent with our policy of limiting Soldier support requirements to those considered essential or non-burdensome.

The military and government civilian labor costs were collected from the Army Military-Civilian Cost System (AMCOS). AMCOS provides complete manpower costs for Active, U.S. Army Reserve, and Army National Guard personnel as well as Army civilians, broken out by grade and occupation. The rates used in these estimates were extracted from the 2005 pay schedules and the FY06 President's budget numbers. They vary by grade but use an all-occupation rate.

We collected contractor labor costs from three different sources. Salary figures for Ph.D. and masters-level industrial-organizational psychologists came from the 2003 Income and Employment Survey published by the Society of Industrial and Organizational Psychology. Hourly rates for test content subject matter experts were constructed after talking with a number of test development companies. The remainder of the labor costs was constructed using Bureau of Labor Statistics data. Specifically, we looked at the 2005 March Supplement to the Current Population Survey and computed the average annual salary for individuals in the occupations of interest. We applied a multiplier of 2.5 to all of the contractor salary costs to reflect the additional costs of overhead and fringe benefits.

Many firms have the expertise to develop and maintain assessment instruments, including multiple-choice test banks, graphics-rich scenario-based tests, and hands-on tests. We spoke to representatives from firms operating in the Washington, DC area that offer test development and maintenance services. We asked them about the steps that go into developing and maintaining a test, the type of resources that would be required, and the staff time/full time equivalents (FTEs) needed per set of test items (for multiple-choice tests), per task (hands-on tests), or per instrument (scenario-based tests).

Given the diversity of firms in this industry, it was important to determine if our resource estimates were representative. Our estimates of labor hours and the hourly contractor rates were benchmarked against additional vendor input obtained through a brief email survey.

Parameters

It was necessary to define fixed values to a number of parameters that affect the cost model. These parameters were:

- Discount rate[1]
- Inflation rate
- Test window
- Personnel costs
- Phase-in period

We set the discount rate at 2% and the inflation rate at 3%. The test window was set at 3 months. Military and government civilian costs are expressed as hourly rates by pay grade. Contractor costs are broken out by job and education level. These detailed personnel costs used are presented in Appendix A. Finally, we assumed a 5-year phase-in period.

Size of the Testing Population

As identified in Chapter 2, testing is assumed for E4, E5, and E6 Soldiers. Because resource requirements depend, in part, on the volume of tests administered, we projected the number of Soldiers who would be in the testing population. We obtained Army personnel authorizations published in June 2005 for FY 2007 through FY 2011. Active, Guard, and Reserve authorizations were each reported by MOS and pay grade. Where necessary, we modified the authorizations to reflect planned changes in the occupational structure (e.g., MOS consolidations). As a result we identified 159 MOS[2] that we consider will comprise the test population domain for the future. Thus, we were able to forecast the average annual population eligible for testing, by component and grade, and, eventually, by the type of tests we forecasted were applicable to the different MOS groupings.

Characteristics of the Cost Model

The cost model consists of a series of interconnected Excel workbooks. The user of the model can specify any of the variable parameters or change any of the assumptions about testing used in the current cost projections. When changes are made in any of these variables, they automatically post to all the other workbooks and outputs. Thus the cost model can continue to function as more refined information is gathered and different testing policy assumptions are made.

---

[1] The *discount rate* is used to calculate the present value of future expenditures or benefits. It is the time value of money, or the rate at which an organization can borrow money to finance current expenditures. If the government can borrow money at a rate of 2%, for example, the present values of an investment of $100 today and $102 one year from now are equal.

[2] In order to simplify the population estimation, we restricted the projection to those MOS that start with entry-level (Skill Level 1) and continue through Skill Level 3. This eliminated some populations such as Special Operations Forces and Recruiter and Retention NCOs.

Activity-Level Cost Estimates

Policy, Oversight, and Coordination Costs

The policy, oversight, and coordination costs are estimated under two scenarios: (a) assuming only Army-wide testing is implemented and (b) assuming that the MOS-testing is fully implemented in addition to the Army-wide testing. Table 3 outlines the staffing assumptions used in the cost estimates.

Table 3. Assessment Office Staffing

	Army-Wide Only	Army-Wide Plus Full MOS Testing
Director/Test Psychologist/GS-15, Ph.D	1	1
Senior Administrative Coordinator/GS-11	1	1
Test Psychologist/GS-13, Ph.D	1	2
Test Analyst/GS-12, MA	1	4
Admin, Contract Specialist /GS-7, GS-9	1	2
Administrative/GS-5	1	5
IT Specialist/GS-9, GS-11	1	1

As described in Chapter 2, the Army Assessment Office performs functions beyond strict overhead activities, including test scheduling, policy setting, and interpretive analysis of test results. It has a critical role in the specification and supervision of contractor support. The estimated annual personnel costs for the office as described in Table 3 range between $630,300 and $1.3 million. The lower bound is the cost for an office assuming only the Army-wide test is implemented. The upper bound is the estimate assuming full implementation of the assessment program. There will be other oversight and coordination costs including travel and IT expenses. The estimated annual travel budget ranges between $12,000 and $30,000.

IT start-up costs will include approximately $50,000 for test development software licensing. This software facilitates development and maintenance of individual test items, item and test banking, and item performance and modification records. There are increasing commercial applications for test development software available. The same firms that offer this capability also generally provide for hosting and delivering the tests, usually on a fee basis per test or per administration. Although normally contracted for as a package, the hosting and delivery costs are addressed later under the test delivery activity.

Under scenario 2 (MOS testing) there will also be a requirement for oversight at each of the MOS proponents. We assumed that all of the MOS test development work will be contracted out; however, as with the Army Assessment Office, there will be a requirement for review and contract administration and guidance. We do not anticipate that this will be a full time requirement at any proponent. Although the costs would vary with the size and activity of the proponent, we estimate that the total in-house (non-contract) costs across the 17 proponents would range from $850,000 to $1.7 million. Most of this cost would be for government civilian personnel.

Assessment Development and Maintenance Costs

There are two primary requirements to factor into the costs of assessment development and maintenance. First, there is the requirement to ramp up a test system including initial job analysis, blueprint development and building item banks. This occurs during the 5-year phase-in period. Once a system is in place, the system must be maintained, but has the advantage of building on the history, data, and test items that have already been developed during phase-in. Assessment development and maintenance costs derive from three main activities: training-oriented occupational analysis, testing-oriented occupational analysis, and test development. The costs are calculated for the Army-wide test as well as for MOS testing.

Chapter 2 described the different types of tests that we postulated to support a very robust and state-of-the-art testing program. The Army-wide tests have an enhanced multiple-choice (EMC) portion and a situational judgment test (SJT) portion. We also conducted an analysis of the 159 MOS identified in our population review and made a decision on how best to test each MOS based on the assessment method(s) (from Table 3) that best matched their performance requirements. We then grouped the MOS by test method(s) building on a strategy that was proposed by Rosenthal, Sager, and Knapp (2005). These groupings, along with our original population estimates, are shown in Table 4. In this projection, all MOS groupings except for one include an EMC test. In addition, some MOS are also tested by other methods as well. The significance of this allocation is in the costing—different assessment methods cost differently to develop and administer.

Table 4. MOS Cluster Descriptions

	Cluster					
	A	B	C	D	E	F
Army-Wide EMC/SJT	●	●	●	●	●	●
MOS Enhanced Multiple Choice	●	●	●	●	●	
MOS Situational Judgment Test		●				
MOS Single Path Scenario			●			
MOS Multiple Path Scenario				●		
MOS Hands-On Tests					●	
Number of MOS	67	10	14	6	14	48
Population (FY07-11 average authorizations across pay grades E4-E6)	74,493	20,094	11,063	2,285	30,671	4,641

The purpose of the MOS clusters was to support the cost analysis based on an assumed distribution of the population by types of tests. The assignment of MOS to test methods is neither definitive nor authoritative. Table 5 shows a sample of the MOS we assigned to each cluster. Although the assignments were done after a review of the MOS characteristics, one could argue for assignment of any MOS to any of the clusters.

Table 5. Sample MOSs by Cluster

Cluster/Test Method	Sample MOSs in Cluster
A. Enhanced Multiple Choice	Field Artillery Surveyor (13S) Cargo Specialist (88H) Public Affairs Specialist (46Q)
B. Situational Judgment Test	Military Police (31B) Chaplain Assistant (56M) Intelligence Analyst (96B)
C. Single Path Scenario	Aviation Operations Specialist (15P) Chemical Operations Specialist (74D) Pharmacy Specialist (68Q)
D. Multiple Path Scenario	Air Traffic Control Operator (15Q) Civil Affairs Specialist (38B) Air Defense C4I TOC Operator (14J)
E. Hands-On Tests	Infantryman (11B) Petroleum Lab Specialist (92L) Cannon Crewmember (13B)
F. No MOS Test (Low Density)	Railway Equipment Repairer (88P) Transmission and Distribution Specialist (21Q) PATRIOT System Repairer (94S)

We should note that one group of MOS (Cluster F) does not have MOS tests under the MOS testing scenario. These are very low density MOS, with total populations in Active and Reserve Components in pay grades E4 through E6 of fewer than 450 Soldiers. From a cost perspective, they have been eliminated from MOS testing. This may not be the eventual policy, but their inclusion cannot be justified from a cost standpoint.

Training-Oriented Occupational Analysis

As discussed in Chapter 2, training-oriented analysis is currently mandated by TRADOC Regulation 350-70. However, this analysis currently is not being performed consistently across MOS. Detailed analysis is a must for support of test development. Therefore, although it is a requirement regardless of whether or not testing is instituted, we have estimated how much it costs to perform this activity. This cost includes the activities of preparing, administering, and analyzing job and task survey data. It also includes the cost of updating and publishing revised source material for task performance such as Soldier's Manuals and other training publications. We obtained data on the cost of training-oriented job analysis through an interview with an official in the Army's Occupational Analysis Office and conducted interviews with TRADOC training analysts for information on the resources required to update task manuals and fully maintain them to support the testing program. We asked about types of labor required, staff time, administrative needs, travel, information technology, and other inputs.

The main cost drivers are contractor labor, differentiated by occupation and skill level and Army civilian labor, differentiated by grade. New investments to enhance information

systems and computer tools are also important. Soldiers and supervisors are an important input into the analysis, because they fill out the job surveys. However, because the Army would pay for the Soldiers anyway, we did not include their time in program costs.[3]

We estimate the cost to perform this analysis for all MOS not in cluster F (Army-wide testing only) to be between $23.5 million and $31.3 million over the program 5-year phase-in period. In accordance with the TRADOC regulation, this analysis should be redone every 3 years at an approximate cost of $141,200 per MOS/skill level combination. We did not estimate the cost of performing an Army-wide (common task) analysis because the Army has been relatively successful at routinely updating this analysis. Because training-oriented analysis is a requirement for the Army to perform anyway and is not being generated solely by the assessment requirement, we do not include any training-oriented occupational analysis cost figures in our subsequent roll-ups for assessment costs.

Testing-Oriented Occupational Analysis

Although training-oriented occupational analysis is useful in test development, there is also a requirement for testing-oriented analysis that specifically addresses test emphasis issues. This information results in the test blueprint. PerformM21 researchers developed a prototype analysis design and associated Soldier survey (Moriarty et al., 2005), but this is a new operational requirement and, unlike the training-oriented analysis, is specific to the testing program.

If the Army did not pursue MOS-specific testing, it would cost approximately $200,000 to develop an automated "testing emphasis" survey to support development and updating of an Army-wide test blueprint. In addition, the Army should budget approximately $25,000 a year for updates and maintenance. If MOS-specific testing is adopted, the software will need to allow for easy adaption to variations in survey content across MOS. It also will need to be used in a variety of locations. We anticipate that the total software costs in this case run approximately $300,000 to develop and $40,000 per year for maintenance and updates.

Testing-oriented analysis needs to be conducted initially to support the phase-in development requirement. Thereafter, testing-oriented analysis normally needs to be re-administered about every 3 years. The labor cost estimate for the initial Army-wide testing-oriented analysis is $103,400 to $137,900 with an annual recurring cost of $25,600 to $34,100. For MOS testing, the initial 5-year cost of testing-oriented analysis is $11.1 million to $14.7 million with recurring annual costs of $2.8 million to $3.8 million. As with the training-oriented analysis, we have not included costs of time for enlisted personnel to complete the surveys.

Army-Wide Test Development

Army-wide test development consists of the contract costs, principally in labor, involved in developing the initial test bank items and test forms for grades E4-E6. The Army-wide test consists of an enhanced multiple-choice test at an initial development cost of $226,700 – $302,300. The range depends on how much overlap in test items exists between the three pay

[3] One could, however, impute an opportunity cost from the value of the activities in which the Soldiers otherwise would have been engaged.

grades. The Army-wide test also has a situational judgment test that we estimate can be developed for $159,200 – $212,300, making the total development costs $385,900 – $514,500.

The annual maintenance costs includes $123,000 – $164,000 for adding test items and managing the enhanced multiple-choice test and $107,000 – $164,000 for adding items and managing the situational judgment test. Total annual maintenance for the Army-wide test is estimated at $230,000 – $328,000.

MOS Test Development

MOS test development and maintenance costs result from two activities: (a) development of enhanced multiple-choice tests, and (b) development of any special tests (see Table 4). We display the total MOS development and maintenance costs by category in Table 6. This table also reflects the phase-in and annual analysis costs. Note that the upper and lower bounds for enhanced multiple-choice tests, situational judgments, and hands-on tests correspond to differing assumptions about the degree of overlap in the content of tests across pay grades. In the case of the path scenarios, the upper and lower bounds are associated with the assumed level of Army subject matter expert support that would be involved in test development[4].

Table 7 lists the initial MOS testing analysis and test development costs broken down by cluster. Each MOS in clusters that include MOS-specific tests (A-E) will receive an enhanced multiple-choice test. MOS in clusters B through E will take an additional type of test. These development costs will be spread out over the 5-year phase-in period.[5]

Table 6. MOS Test Development and Maintenance Costs

Activity	Cost Estimate
Phase-In	
Testing-oriented analysis	$11.0M - $14.7M
Enhanced Multiple Choice	$25.2M – $33.6M
Other tests[a]	$13.5M – $14.8M
Total Development Cost	*$50.0M - $62.6M over phase-in period*
Annual	
Testing-oriented analysis	$2.8M - $3.8M
Enhanced Multiple Choice	$13.7M – $18.2M
Other tests[a]	$1.2M - $1.6M
Annual Maintenance Cost	*$17.7M - $23.6M*

[a]Includes situational judgment, path scenario, and hands-on tests.

[4] Whereas we assume that the other tests would be developed by contractors with subject matter expertise provided primarily by former Army personnel, we assume development of the scenario tests will require direct involvement of a small cadre of active duty (or reserve component) personnel.

[5] Tables 7 and 8 reflect a different treatment of the hands-on tests by breaking them into two categories based on the characteristics of the MOS being tested with this method. Hands-on model E1 assumes to have from 2-4 tasks tested hands-on while model E2 has about 10 tasks tested each year. Model E2 would require building a bank of tasks to be tested while the E1 model tasks remain more static.

Table 7. MOS Initial Test Development Costs: Breakdown by Cluster

MOS Cluster	# MOS	Testing-Oriented Analysis	Enhanced Multiple-Choice	Other Test	Total
A (EMC only)	67	$6.7M - $8.9M	$15.2M – $20.3M		$21.9M – $29.1M
B (EMC + SJT)	10	$1.0M - $1.3M	$2.3M – $3.0M	$1.6M – $2.1M	$4.9M – $6.5M
C (EMC + SPS)	14	$1.4M - $1.9M	$3.2M – $4.2M	$6.3M – $6.9M	$11.4M – $12.9M
D (EMC + MPS)	6	$0.6M - $0.8M	$1.4M – $1.8M	$3.4M – $3.6M	$5.6M – $6.2M
E1 (EMC + HO)	9	$0.9M - $1.2M	$2.0M – $2.7M	$1.1M	$4.0M – $5.0M
E2 (EMC + HO)	5	$0.5M - $0.7M	$1.1M – $1.5M	$1.1M	$2.7M – $3.2M

Note. EMC = Enhanced Multiple Choice, SJT = Situational Judgment Test, SPS = Single Path Scenario, MPS = Multiple Path Scenario, HO = Hands-on Test.

Table 8 breaks out the annual maintenance costs for analysis and development by cluster. Clusters C and D do not include annual developmental maintenance costs for the path scenario tests because these tests do not require routine maintenance. Rather, they only get updated when the equipment or doctrine on which the Soldiers are being tested changes.

Table 8. MOS Annual Test Maintenance Costs: Breakdown by Cluster

MOS Cluster	# MOS	Testing-Oriented Analysis	Enhanced Multiple-Choice	Other Test	Total
A (EMC only)	67	$1.7M - $2.3M	$8.2M – $11.0M		$9.9M – $13.3M
B (EMC + SJT)	10	$0.2M - $0.3M	$1.2M – $1.6M	$1.1M – $1.4M	$2.6M – $3.4M
C (EMC + SPS)	14	$0.4M - $0.5M	$1.7M – $2.3M		$2.4M – $2.8M
D (EMC + MPS)	6	$0.1M - $0.2M	$0.7M – $1.0M		$0.8M – $1.2M
E1 (EMC + HO)	9	$0.2M - $0.3M	$1.1M – $1.5M	$0.6M	$1.9M – $2.4M
E2 (EMC + HO)	5	$0.1M - $0.2M	$0.6M – $0.8M	$0.8M	$1.5M – $1.8M

Note. EMC = Enhanced Multiple Choice, SJT = Situational Judgment Test, SPS = Single Path Scenario, MPS = Multiple Path Scenario, HO = Hands-on Test.

Assessment Delivery Costs

Although all the tests (except for hands-on) will be computer based and delivered, there are still some delivery costs that must be calculated. As outlined in Chapter 2, we assume that test delivery will be through DTF/DTTP facilities that the Army currently owns. From our interviews with the DTF managers, it appears the DTFs have more than sufficient excess capacity to house the entire testing program at no additional cost to the Army.[6] The problem is that there are large geographic areas that are not served by a DTF. The National Guard has a similar set of distance learning facilities—DTTPs. They also have the ability within their armories to allow web-based testing of unit members. The seat capacity and geographic coverage

[6] COR note: excess capacity was the situation in 2006; there are indications it may no longer be true in 2011.

provided by all of these options appears to be more than adequate for an assessment program of this size. We were unable to determine if there would be any additional costs to the Army of using the DTTPs or armory facilities. We also considered the possibility of using commercial facilities if the DTF/DTTP coverage was inadequate. We received a quote from one of the largest computer-based test delivery companies of $35 - $73 per test to use their facilities for test delivery. The range depends on the length and type of test administered. Because of the uncertainty of this requirement, we have not included this option in our costing.

Although the cost of the DTF/DTTP is negligible, there is still a requirement to host and deliver the tests. Again, we assumed the Army would contract this function, and we identified an annual per assessment/per examinee cost of $10 for this service through commercial sources. Depending on the Army's ability to negotiate a rate discount, we estimate the total annual licensing fee to range between $1.4 million and $3.2 million.

Because the testing is high stakes testing, there is a requirement that the tests be proctored to verify Soldier identification and to monitor test taking. It could be argued that this is a legitimate supervisor responsibility, in which case the cost would be zero. However, we have figured in the costs of contract proctors, in case military proctors are not feasible. If the Army wanted to maximize scheduling flexibility, they could hire full-time proctors to staff the DTFs during the entire period they are currently open. The cost of these contract proctors could amount to $8.9 million. For Army-wide testing only, we estimate that at most the Army would need to provide quarter-time proctors for a total maximum cost of $2.2 million.

The final area of test delivery costs is in the hands-on tests. We believe the most effective way to administer these assessments would be to have contractor scoring teams that travel from location to location administering and scoring the hands-on tests. We estimated the costs for these teams to be $11.5 million.

Soldier Costs

The final area of costing is the time of the Soldiers who are being tested. Since these Soldiers are being paid regardless, this is more an opportunity cost than a budget cost. It could be argued that since testing will improve selection and Soldier readiness, test time is as productive as other training functions Soldiers perform. Nonetheless, test time is time that Soldiers cannot be performing other functions, so it can also be looked on as a cost. We estimate the cost of Soldier time if the entire program is implemented to be $25 million. The cost if just the Army-wide assessment is implemented will be around $9.6 million.

Summary of Cost Estimates

Figure 5 displays the total Army-wide exam costs for test development, maintenance, and Soldier costs. Training-oriented occupational analysis survey work and costs associated with the possibility of test administration at commercial computer-based test centers are not included.

Figure 6 displays the same information for the entire program including MOS-testing. The upper and lower bounds reflect the different assumptions we have made regarding overlap of

pay grades and software licensing fees. Note that the annual maintenance cost numbers include annual maintenance of the tests, delivery costs, and oversight costs.

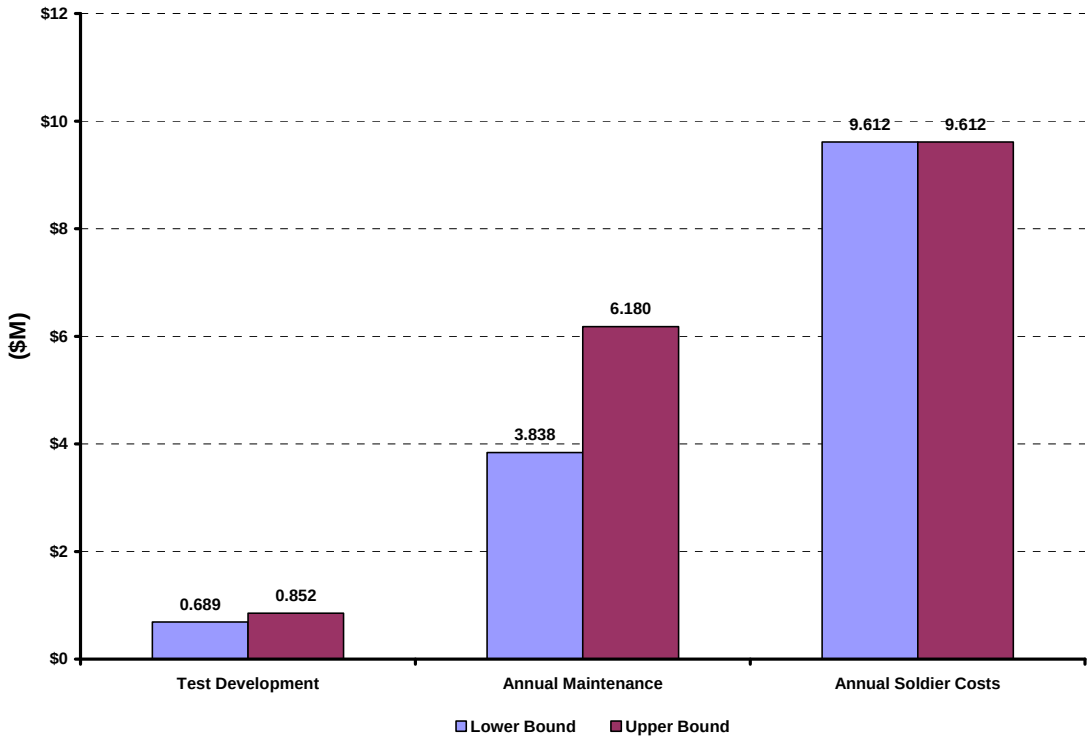

Figure 5. Total Army-wide exam costs.

For the Army-wide assessment, the cost that represents the value of the time Soldiers spend testing is more than twice the annual maintenance costs of the assessment program. The initial development cost is quite small, in part because of the work that has already been done in the test development portion of the PerformM21 project to prepare for an Army-wide exam.

For the MOS testing, costs are quite high because of the large number of different tests that need to be developed and the inclusion of several high cost test methods for some MOS under the notional program. Spread over a phase-in period of 5 years, the annual development costs would be between $7.4 million and $9.6 million. The annual maintenance costs are also quite a bit larger if the complete assessment program is implemented, though $11.5 million of that is the cost for the hands-on scoring teams.

Finally, Figure 7 estimates costs *for a single MOS* within each cluster that includes MOS-specific testing (A-E). To simplify this presentation, we show the average of the lower and upper bound estimates presented in previous figures. Army-wide assessment costs are not included in these estimates. This figure illustrates that Cluster A (enhanced multiple choice tests only) costs, taken across development, maintenance, and Soldier costs, are considerably lower than for the clusters that include additional test methods.

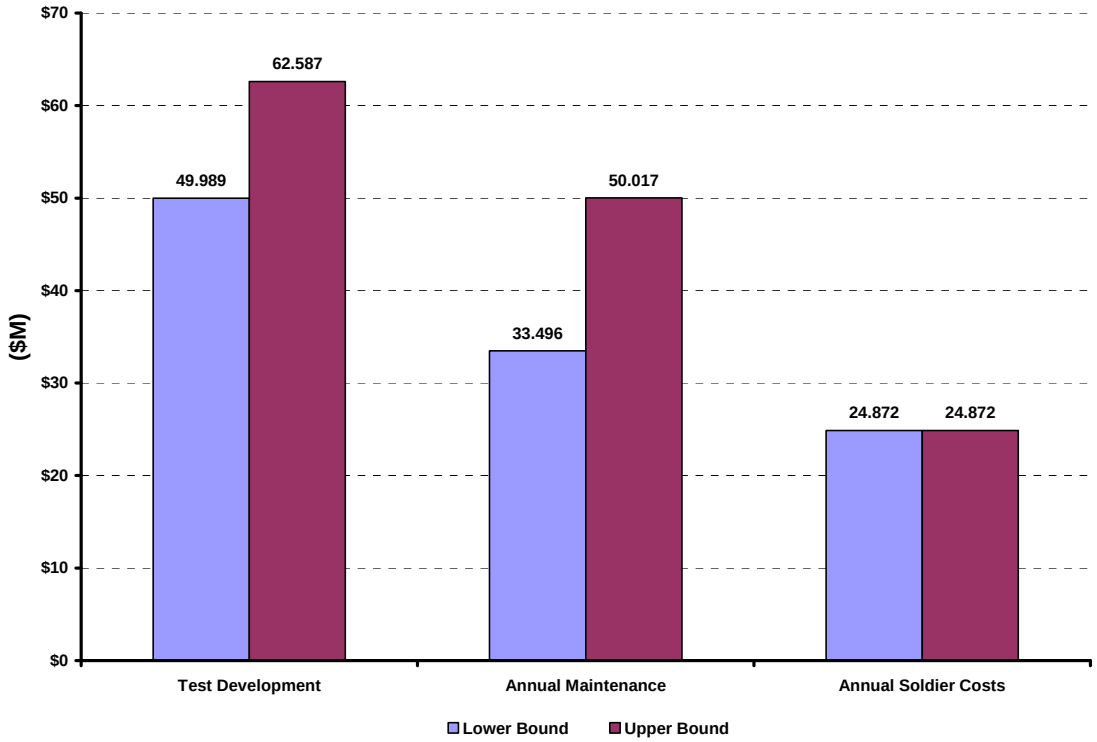

Figure 6. Total program costs including Army-wide and MOS testing.

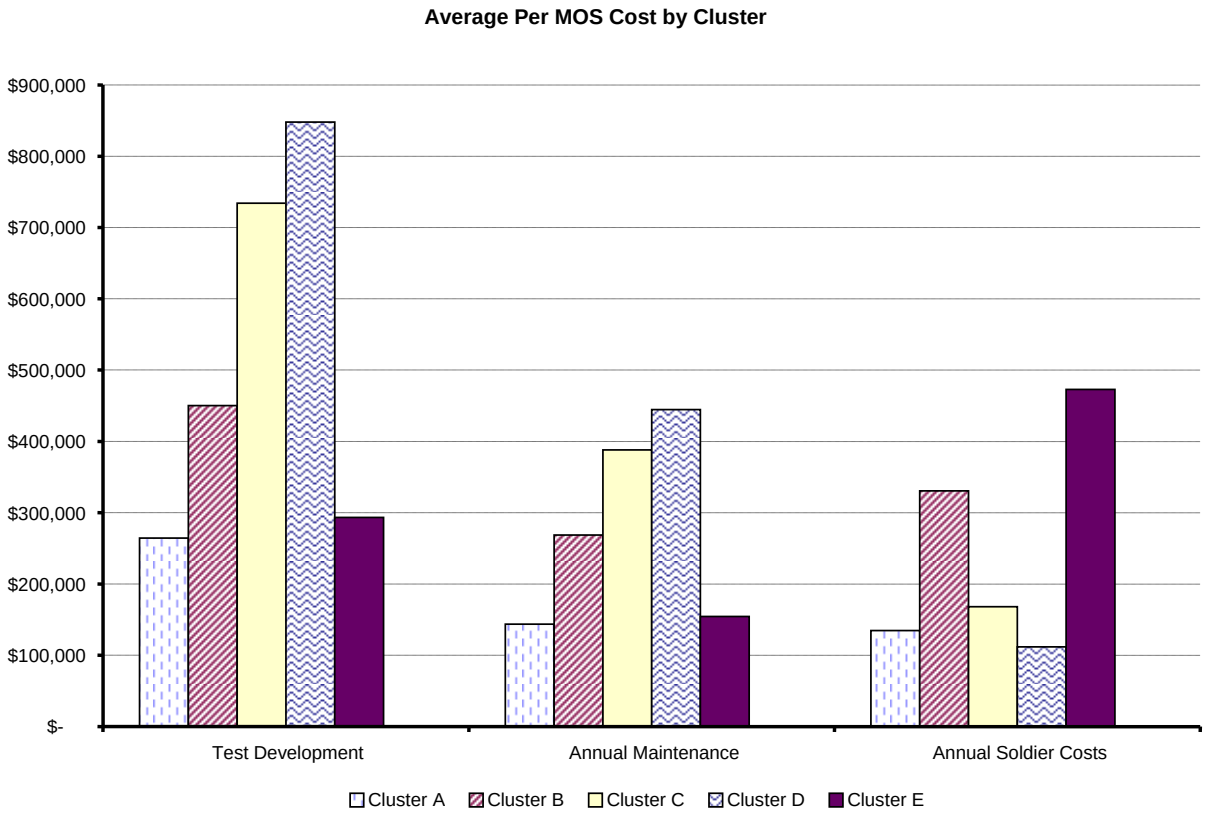

Figure 7. Average per-MOS cost of MOS-specific testing in each MOS cluster.

Benefits

The purpose of the PerformM21 project was to design a test program that would be both effective and affordable, which would in turn support enlisted personnel selection and classification research. The question of affordability is addressed with the cost analysis just described. Here we look at the ways in which a test program could be viewed as being effective for addressing important Army needs.

The assessment program would improve overall readiness and performance in a number of ways. Foremost is the effect the program should have on Soldier performance and readiness. It also would provide better information on personnel readiness to Army leadership, thereby allowing for more informed decisions. Additionally, there are a number of research applications of the assessment data that would further enable the Army to select, retain, and motivate high-quality Soldiers. Figure 8 illustrates the paths by which the Soldier assessment program can lead to a more capable force. It can lead to improvements in the promotion process, in leadership evaluation of readiness, and in the selection and classification research arena.

Benefits Through Improved Soldier Readiness and Performance

The ultimate objective of any Army investment related to personnel is a more capable force. The proposed assessment program could achieve this objective both by improving the selection of Soldiers for promotion and by motivating all Soldiers interested in an Army career to pursue self-development and achieve a higher level of competence in the knowledge and skills required for their jobs.

In practice, these gains in productivity, readiness, and performance could be realized in two ways. First, the Army could reduce endstrength (perform the same mission with fewer, albeit more capable, personnel). Alternatively, the Army could accomplish more with the current endstrength. Given the current demands on the Army, the latter option is more likely.

The assessment program would promote a more capable force by acting as a selection tool and a motivational tool. First, the Army could use the results of the assessment program to help identify the most capable Soldiers and move them into leadership positions. Second, making the assessment program an integral part of the advancement process will motivate Soldiers to work harder and expand their capabilities. Additionally, training may become more efficient as the assessment program reveals which areas require traditional training emphasis and which areas are most appropriately addressed through Soldier self-development.

These benefits depend on the motivation of Soldiers to take the assessment process seriously. If the assessment program is not tied to some aspect of the Soldier's career (e.g., promotion opportunities) as was the intention at the start of the PerformM21 effort, he or she will not have a vested interest in performing well on the assessment. The benefits depend as well on a direct relationship between the actual knowledge and skills required for a Soldier's job and the items tested in the assessment process. In other words, preparation for the assessment should actually make Soldiers more job-capable, not just prepare them for testing.

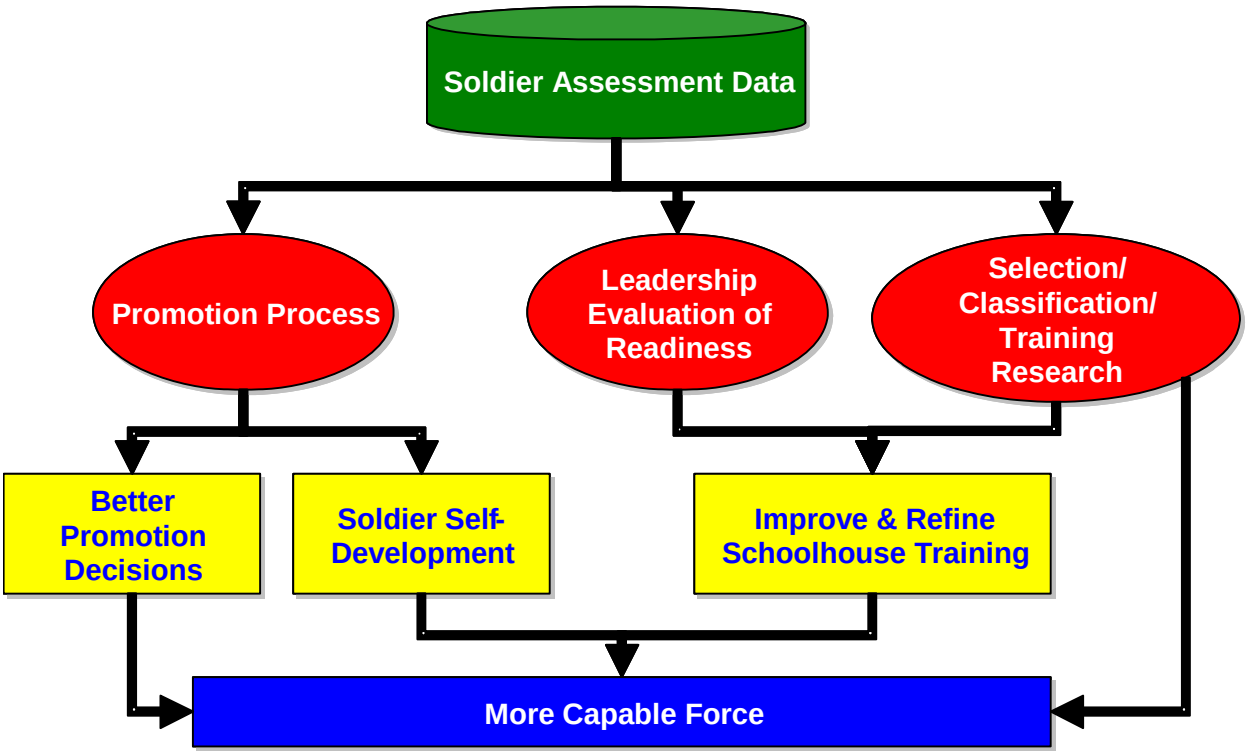

Figure 8. Sources of benefits from Soldier assessment data.

Benefits from Better Information on Personnel Readiness

The primary benefit of an assessment program is its effect on Soldier capabilities. However, the availability of better information on capabilities, readiness and performance is, in and of itself, valuable to Army leadership. Assessment data—beyond its use in the promotion process as described above – can help Army leaders at all levels make more informed resource-allocation and personnel-management decisions.

Information on individual readiness could provide additional information on unit readiness. Leaders could also make more informed training and assignment decisions. The assessment data could both highlight areas of weakness in need of additional investments and show where traditional training could be streamlined. This could significantly improve the efficiency of training, allowing the Army to focus on weaknesses both at the unit and individual level.

Benefits from Research Applications of Assessment Data

The Army and the other Services must conduct selection and classification research to determine how best to use scores from the Armed Services Vocational Aptitude Battery (ASVAB) and to develop new and better predictors. The purpose of this area of research is to more accurately identify, recruit and distribute individuals into jobs for which they are well suited. The selection process involves the identification of qualified recruits, while the classification process directs recruits into jobs for which they are qualified.

One critical part of this type of research is the development of job performance criterion measures. In other words, the Army needs to identify the critical knowledge, skills, and attributes associated with each MOS so that it can identify the candidates most likely to have the aptitude to acquire those capabilities. Developing such measures for a purely research application would be very expensive. However, the Army-wide tests can provide useful criteria for selection research and the MOS tests can provide criterion measures for classification research.

Test scores from the program may also provide useful information for other types of research (e.g., training evaluation). The tests can offer a "before and after" view of Soldier capabilities that would allow researchers to measure the impact of various investments on the quality and performance of the force.

Contrast with Prior Test Program

In 1982, the Government Accounting Office (GAO) published a cost-effectiveness evaluation of the Army's SQT program (GAO, 1982). GAO's evaluation framework was similar to that used in this analysis in that the SQT program's costs were expressed in dollar terms, and compared to qualitative assessments of the benefits realized from the program.

Costs were difficult to compute because "no one in the Army has any complete record of actual costs to develop the SQTs" (GAO, 1982, p. 12). The costs of test development were extrapolated from Army schools' budget data. The costs of printing the tests and managing the program were collected from the Army Training and Support Center. GAO gathered case study evidence of the time and resources that units required to organize and administer the tests.

According to the GAO report, more than $25 million was spent annually to develop, print, distribute, and score hundreds of skill qualification tests. This cost estimate—which equates to $55 million in 2005 dollars—was a lower bound because the costs of test administration were not included and because many MOS had yet to take part in the program. Although GAO was not able to express the burden to units of organizing and administering the SQT in dollar terms, the evidence was that time, equipment, and material costs were significant. In part, this is because each MOS test included hands-on and job-site components in addition to a written multiple-choice test. Table 9 shows the cost breakdown presented by GAO for FY 1981, with amounts expressed in 2005 dollars. The test development costs include military labor, overhead, and Soldiers' time to validate the tests.

Table 9. Costs of SQT Program from GAO Evaluation

Expense Element	Costs in FY 1981 (Millions $2005)
Test Development	37.818
Printing of Tests	6.135
Program Management	11.175
Subtotal	55.128
Test Delivery/Administration	Unknown
Soldier Time	Unknown

These costs are difficult to compare to the estimates developed in the PerformM21 analysis for several reasons. The SQT at the time of the GAO report included only an MOS component and some MOS had not yet become part of the SQT system. Printing costs are unlikely to be a significant cost driver today. And although the GAO did not provide a dollar estimate of administration costs, we expect that they will be lower under the new program due to greater emphasis on computer-based multiple choice testing relative to hands-on and job-site testing.

Summary

The cost-benefit analysis of the PerformM21 assessment program demonstrates that it would represent a significant investment of resources for the Army. Under an extreme set of assumptions (full MOS testing, full range testing), test-development costs could total nearly $60 million over the phase-in period; annual maintenance costs could approach $53 million. On top of these costs, the testing process would require almost $25 million annually in Soldier time. An option for only Army-wide testing stands in marked contrast to this, with initial test development costs for phase-in at $515,000 with annual maintenance costs at about $4.4 million and Soldier costs of under $10 million. Because the testing program has not yet been well defined, however, the total costs for either approach could vary in either direction quite substantially.

These significant program costs should be balanced against the potential for substantial benefits. The main effects of the testing program will be to make the force more capable. It can accomplish this goal both through better promotion selection and by providing increased incentives for Soldiers to improve through self-development. In addition, the program will give the Army a better source of data on Soldier readiness, capabilities and skills; this information will provide for better decisions about resource allocations and could help to streamline traditional training pipelines.

When the Army must accomplish expanding missions with fixed or diminishing numbers of Soldiers, proactive investments are critical. The benefits enumerated here are only possible if the tests are designed to encourage Soldiers to improve in all areas critical to their jobs, and if the Soldiers perceive a relationship between performing well on the test and succeeding in their Army careers.

CHAPTER 4: SUMMARY AND DISCUSSION

Roy C. Campbell (HumRRO)

The PerformM21 research effort started in December 2002 and covered three distinct but related phases, with the final phase (including this report) being completed in early 2006. While we accomplished the technical objectives outlined in each of the phases, we also learned a great deal about implicit requirements and consequent effects often not generally foreseen in the technical objectives. Moreover, the period of the PerformM21 effort has been a tumultuous one for the Army; we recognize this as a much different Army than when the project work started in late 2002. Fortunately, though the involvement of the ATPAT, we have been able to keep the PerformM21 effort on track, while still incorporating and reacting to cultural, operational, and priority fluctuations within the Army.

The scope of the PerformM21 work over the three phases was too broad to comprehensively summarize in this chapter. However, we do present our primary conclusions and we will also address and discuss some of the less obvious, but perhaps more imperative, aspects of an Army test program as we look back on what we have learned over the life of the project.

Major Conclusions

Army-Wide Testing

The main conclusion from the PerformM21 work is that Army-wide testing is not only feasible but that it is cost effective as well. While the cost figures in this report cover a range of options, and are not final, they nonetheless are within the realm of what could be considered "reasonable," especially considering the very quick payback in terms of enhancement of selection for promotion and reinforcement to the self-development program. Moreover, PerformM21 has done much of the preliminary work, including initial blueprint work and substantial item development.

MOS Testing

MOS testing is a more complex issue. Our conclusion, however, is that it is just as feasible as Army-wide testing, although the cost issue is less resolved or apparent. It must be pointed out that we investigated a full range of innovative job testing options, some of which are quite expensive to develop and/or administer. The Army is not bound by those options, and less expensive, but still effective, combinations of test methods could be used instead (e.g., just enhanced multiple choice and situational judgment tests). It also is time to do away with the concept that all MOS must have the same programs; in this case, that everyone must test. In reality, some MOS are ready for almost immediate testing and others will probably never test. This is neither inconsistent nor unfair. Soldiers are promoted by MOS and grade and do not compete with Soldiers in other MOS or grades. Many MOS already have different requirements, or have different opportunities to earn promotion credits, that other MOS do not have. Each

Branch Chief must decide the best way to select and skill-qualify Soldiers in their Branch—for some this will include direct competency assessment and others will not select this route.

Selected Discussion Points

Integration into the NCO Development and Promotion System

The work done in assessment and the development of an assessment delivery system under PerformM21 is only part of a much larger picture. NCO development for the 21st century is a many-faceted program encompassing changes in NCO selection, education, self-development, utilization, and assignment. Within TRADOC, the entire NCO leader development and accountability program is undergoing major review and revision. The work done previously under ARI's *Maximizing 21st Century Noncommissioned Officer Performance* (NCO21) project (Knapp, Heffner, & Campbell, 2003) lays the groundwork for an overall transformation of the semi-centralized promotion system, of which the PerformM21 assessments would be a part. These related, but independent efforts, within ARI and elsewhere, need to be melded into a coherent system to serve the Army and NCO corps in the coming years.

Buy-In to Assessment

It is important to recognize that a return to routine testing will be a culture change for the Army. Moreover, it is not just individual Soldiers who need to be convinced the assessment program is worth doing and doing well. The more that all supporting organizations and individuals are vested in and respect the idea, the more effective it will be. It will therefore be important to "market" the program to all stakeholders by informing them of program goals and plans while also educating them about the various considerations and constraints that go into developing a successful program. People invariably underestimate what it takes to develop and maintain a high quality assessment program and they often question the motives of those developing those programs.

Commitment to Quality

If the Army is serious about instituting an assessment program, it must be willing to do so without cutting corners in the quality of that program. This is not only a cost issue; it is the dedication to adhere to the proven tenets for development and administration of tests. For as important as it is to have a testing system in place, a low-quality assessment program would be worse than none at all. A poorly administered, inadequately maintained, impermanent program would be unfair to Soldiers and would not help the Army improve readiness. Commitment to quality is critical to the program's success.

Flexibility in the Assessment Planning

As indicated, we are well aware that the conditions in the Army today are not the same as they were when we started the PerformM21 work more than 3 years ago. And while we firmly believe the needs that gave impetus to the competency assessment movement, including the findings from the NCO ATLDP, are as valid now as they were then, we realize that operational

considerations and the priorities of the Army have changed. So even if 2006 is not the right time for a promotion assessment program, the right time will re-emerge. Meanwhile, there are many ways to take advantage of the ideas and progress that have been made under the PerformM21 work. The current plan is to proceed with a feasibility investigation of limited testing for select MOS to support reclassification decisions. This will build off the work performed to date. As long as such a program meets some minimal criteria for test preservation and security, it will reinforce the long-term goal of eventual wide-spread competency assessment.

Conclusion

The PerformM21 project has made too much progress and has produced too many usable products to be consigned to the category of "just another report." Nonetheless, it is still very much a work in progress because many of the assumptions that went into the program design are subject to Army interpretation and operational examination before adoption. This is especially true of the cost and benefits analysis presentation—the cost model is purposely flexible in that it can accommodate changes in many of the variables and assumptions used in the presentation. But there is an element of currency in all research and more so, perhaps, in the PerformM21 work. Much of the development work and many of the products are pervious to the passage of time. The Army planning process for an effective competency assessment program should not be shortcut. The sooner that internal Army process starts, the more relevant is the current work.

REFERENCES

Department of the Army. (2002). *The Army training and leader development panel report (NCO).* Final Report. Fort Leavenworth, KS,: U.S. Army Combined Arms Center and Fort Leavenworth.

Keenan, P.A., & Campbell, R.C. (2006). *Development of a prototype self-assessment program in support of soldier competency assessment* (Study Report 2006-01). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D.J., & Campbell, R.C. (Eds.) (2004). *Army enlisted personnel competency assessment program: phase I needs analysis* (Technical Report 1151). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D.J., & Campbell, R.C. (Eds.) (2005). *Army enlisted personnel competency assessment program: phase II report* (Technical Report 1174). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D.J., Heffner, T.S., & Campbell, R.C. (2003). *Recommendations for an Army NCO semi-centralized promotion system for the 21st century* (Research Report 1807). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Moriarty, K.O., & Knapp, D.J. (Eds.) (2007). *Army enlisted personnel competency assessment program: phase III pilot tests* (Technical Report 1198). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Moriarty, K.O., Knapp, D.J., & Campbell, R.C. (2006*). Incorporating lessons learned into the Army competency assessment prototype* (Study Report 2006-08). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Rosenthal, D., Sager, C.E., & Knapp, D.J. (2005). *A strategy to produce realistic,cost-effective measures of job performance* (Study Note 2005-03). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

United States General Accounting Office (1982). *The Army needs to modify its system for measuring individual soldier proficiency* (Report to the Secretary of the Army, FPCD-82-28), March 30.

APPENDIX A
DETAILED PERSONNEL COST ASSUMPTIONS

This appendix provides more detailed information on the types, quantities, and costs of personnel required for each major activity.

Test Development and Maintenance Labor Categories

We anticipate that three major categories of personnel would be bid to work on test development and maintenance:

1. Ph.D.-level Testing Psychologist
2. Masters-level Testing Psychologist
3. Subject Matter Experts (SMEs)

The Ph.D.-level testing psychologists are responsible for psychometric decisions, technical and management oversight, and conducting/overseeing data analysis (e.g., item analysis, reliability estimates).

The Masters-level psychologists train SMEs and facilitate SME meetings to review/revise test items. They also edit draft test items and provide feedback to SME item writers. Other tasks in this labor category include:

- Managing the item bank
- Selecting items for multiple forms
- Conducting data analysis
- Generating preliminary and final scores
- Developing and maintaining examinee study guides

SMEs will most likely be retired military personnel. They will participate in item writing and reviewing training; identify reference sources, graphics, and other materials to support item development; write and review test items; and review test forms to ensure non-redundant content.

Test Development Labor Hour Estimates

Table A1. Labor Hour Estimates for Development of E4 EMC Test

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	100	$128.61
Masters	1	200	$91.35
SME	4	400	$75.00

Table A2. Labor Hour Estimates for Development of E4 SJT Test

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	100	$128.61
Masters	1	200	$91.35
SME	4	250	$75.00

Table A3. Labor Hour Estimates for Development of E4 Single Path Scenario

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	720	$128.61
Graphics Artist	1	320	$50.63
Multimedia Developer	1	320	$78.19
Production Manager	1	720	$75.34
Computer Programmer	1	160	$80.00
Army SME	2	368	$51.34

Table A4. Labor Hour Estimates for Development of E4 Multiple Path Scenario

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	720	$128.61
Graphics Artist	1	320	$50.63
Multimedia Developer	1	480	$78.19
Production Manager	1	1,080	$75.34
Comp. Programmer	1	320	$80.00
Army SME	2	368	$51.34

Table A5. Labor Hour Estimates for Development of E4 Hands-On Test, Model 1

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	8	$128.61
Masters	1	60	$91.35
SME	2	40	$75.00
Scorers	2	8	$75.00

Table A6. Labor Hour Estimates for Development of Hands-On Test, Model 2

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	32	$128.61
Masters	1	168	$91.35
SME	6	48	$75.00
Scorers	2	16	$75.00

Test Maintenance Labor Hour Estimates

Table A7. Labor Hour Estimates for Maintenance of E4 Level EMC Test

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	90	$128.61
Masters	1	180	$91.35
SME	4	180	$75.00

Table A8. Labor Hour Estimates for Maintenance of E4 SJT Test

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	90	$128.61
Masters	1	180	$91.35
SME	4	145	$75.00

Table A9. Labor Hour Estimates for Maintenance of E4 Hands-On Test, Model 1

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	8	$128.61
Masters	1	24	$91.35
SME	1	40	$75.00

Table A10. Labor Hour Estimates for Development of Hands-On Test, Model 2

	Number	Hours per Employee	Hourly Rate
Ph.D.	1	8	$128.61
Masters	1	56	$91.35
SME	2	60	$75.00

Table A11. Labor Costs by Category

Labor Costs	Annual Salary	Hourly Rate
Contractor Costs		
Ph.D	$107,000	$128.61
Masters	$76,000	$91.35
SME	$62,400	$75.00
Graphics Artist	$42,120	$50.63
Multimedia Dev	$65,050	$78.19
Production Manager	$62,687	$75.34
Computer Programmer	$66,560	$80.00
Admin Support	$35,360	$42.50
Active Army Personnel		
E4	$62,944	$30.26
E5	$78,084	$37.54
E6	$92,299	$44.37
E7	$106,787	$51.34
E8	$117,761	$56.62
E9	$143,010	$68.75
Reserve Army Personnel		
E4	$9,254	$30.44
E5	$9,648	$31.74
E6	$10,183	$33.50
National Guard Personnel		
E4	$9,056	$29.79
E5	$9,659	$31.77
E6	$10,234	$33.66
Army Civilian Personnel		
GS-5	$43,761	$21.04
GS-7	$53,883	$25.91
GS-9	$66,705	$32.07
GS-11	$81,877	$39.36
GS-12	$98,565	$47.39
GS-13	$118,028	$56.74
GS-15	$167,469	$80.51